

# Research Training for Releasable Animals

GORDON B. BAUER

Division of Social Sciences, New College of Florida, 5700 North Tamiami Trail, Sarasota, FL 34243, U.S.A.,  
and Mote Marine Laboratory, Sarasota, FL 34236, U.S.A., email bauer@ncf.edu

---

**Abstract:** *Restrictions on training potentially releasable animals such as those undergoing rehabilitation care or wild-caught captives have limited our understanding of sensory processes, cognition, and physiology important for conservation of species. It is common practice among several U.S. federal agencies to limit training of animals available for release. The behavioral argument justifying this practice is that training habituates subjects to people and conditions them to associate people with rewards such as food; habituation to and positive associations with people will lead animals into dangerous situations after their release. If under special circumstances research training is permitted, all trained behaviors must be extinguished before release because behaviors will transfer to the natural setting. Research on animal learning and memory indicates that these may not be accurate scenarios. A review of the literature on habituation, classical and instrumental conditioning, and compound conditioning suggests that learning within a research setting does not add to learning that already occurs in procedures associated with basic feeding and care. In fact, animals probably learn less about people in a training setting. Furthermore, context-specific effects on memory limit behavior transfer from captive to natural settings. Extinction is strongly susceptible to context effects, which suggests that extinction does not effectively transfer to the postrelease setting. Counterintuitively, extinction of responses to experimental stimuli under some circumstances may enhance undesirable learning about humans. Under those circumstances in which isolation from human contact is difficult or undesirable, behavioral research can present an ideal format for minimizing learning about humans and provide biological information important for conservation.*

**Key Words:** animal learning, animal memory, animal release, policy

Investigación para el Entrenamiento de Animales Liberables

**Resumen:** *Las restricciones para el entrenamiento de animales potencialmente liberables, como los que están en cuidado de rehabilitación o criados en cautiverio, han limitado nuestro entendimiento de procesos sensoriales, cognición y fisiología importantes para la conservación de especies. La limitación del entrenamiento de animales disponible para liberación es una práctica común en varias agencias federales de E.U.A. El argumento conductual que justifica a esta práctica es que el entrenamiento habitúa a los sujetos a personas y los condiciona a asociar personas con recompensas, como alimento; la habituación a y las asociaciones con personas conducirá a los animales a situaciones de peligro después de su liberación. Si se permite el entrenamiento bajo circunstancias especiales, todas las conductas entrenadas deberán extinguirse antes de la liberación porque las conductas serán transferidas al medio natural. La investigación sobre el aprendizaje y memoria animal indica que estos pueden ser escenarios incorrectos. La revisión de literatura sobre habituación, condicionamiento clásico e instrumental y condicionamiento compuesto sugiere que el aprendizaje en un ambiente de investigación no se agrega al aprendizaje que ocurre en procedimientos asociados con alimentación y cuidado básicos. De hecho, los animales probablemente aprenden menos sobre personas en un ambiente de entrenamiento. Más aun, la transferencia de conducta de ambientes de cautiverio a naturales está limitada por efectos de contexto específico sobre la memoria. La extinción es altamente susceptible a los efectos de contexto, lo que sugiere que la extinción no se transfiere efectivamente al ambiente posterior a la liberación. Contraintuitivamente, la extinción de respuestas a estímulos experimentales bajo algunas circunstancias puede reforzar el aprendizaje sobre humanos no deseado. Bajo esas circunstancias en las que*

*el aislamiento del contacto humano es difícil o indeseable, la investigación sobre conducta puede presentar un formato ideal para minimizar el aprendizaje sobre humanos y proporcionar información biológica importante para la conservación.*

**Palabras Clave:** aprendizaje animal, liberación de animales, memoria animal, política

---

## Introduction

Animal regulatory agencies in the United States restrict behavioral research on many captive, releasable species. Although pre- and postrelease training for purposes of reintroduction (Kleiman 1989) or veterinary care may be permitted, training for basic biological research is frequently not. For example, National Oceanic and Atmospheric Administration (NOAA) regulations (2003) and guidelines for release of stranded marine mammals including cetaceans, pinnipeds, otters, and manatees (U.S. National Marine Fisheries Service [NMFS] and U.S. Fish and Wildlife Service [USFWS] 1997) discourage human interactions, including the training necessary for many types of research with captive, releasable animals. A NOAA regulation (50 CFR 216.27) states that "marine mammals undergoing rehabilitation or pending disposition... shall not be trained for performance..." The NMFS and USFWS guidelines for release (1997: 38) state, "In order to prevent the acquisition of unnatural behaviors, interactions with humans should be kept to a minimum, and limited to such activities as force-feedings, treatments, etc."

The behavioral justifications for minimizing contact and training may be summarized as follows: Humans constitute a major threat to animals in their natural habitat, for example, through provisioning with inappropriate foods, death and injuries from boat strikes, death in fishing nets, and willful killing. If animals are habituated to humans in captive settings and associate humans with rewards, they will be likely to approach or at least not actively avoid humans in natural settings. Attraction to humans or failure to avoid them in the wild is ultimately a threat to animal health and survival. Because experimental, behavioral research in captive settings involves close contact between humans and animals, it should be discouraged.

Restrictions on behavioral experimentation have serious consequences because they minimize opportunities for studies on animal sensory processes, cognition, behavior, and physiology which in turn limit development of important knowledge necessary for protecting animals in the wild. For example, the Florida Manatee Recovery Plan (U.S. Fish and Wildlife Service 2001) identifies objectives that require laboratory studies for thorough explication. Objectives such as minimizing deaths due to boat strikes and water control structures require the careful analysis of sensory processes such as hearing and touch that only controlled study in a laboratory can provide. Studies demanding frequent measurement from captive manatees

trained to provide blood and urine several times a week allowed Manire and colleagues (2003) to model some of the physiological effects of release, another recovery-plan objective. More such studies are needed.

Several recent reports suggest an absence of transfer of trained behavior from captivity to natural settings, a finding inconsistent with the need for restrictions on animal training. Gales and Waples (1993) and Wells et al. (1998) both report that released bottlenose dolphins (*Tursiops truncatus*) did not demonstrate behavioral transfer despite extensive training in captivity. In the former example, behaviors explicitly trained in captivity for use in the wild were not expressed after release. Similarly, Fellner et al. (2005) report that manatees failed to exhibit behaviors trained in captivity after they had been released.

The justification for minimizing behavioral experimentation with releasable animals is based on hypotheses that have not been tested empirically. They would be difficult to test because of the problem of implementing the appropriate factorial experimental design and establishing baseline levels of relevant behavior of appropriate control groups in natural settings. The hypotheses can, however, be evaluated through consideration of the laboratory-based experimental literature that addresses how animals learn and remember. Although studies of rats, pigeons, and to a lesser extent rabbits are most frequently reported in this literature, the rules of learning show considerable generality across both invertebrates and vertebrates (reviews in Macphail 1982; Pearce 1997; Papini 2002; Domjan 2003). The diverse aspects of learning have not been comprehensively studied comparatively across all species, but the similarities of learning phenotypes that have been studied are striking (Macphail 1982; Papini 2002).

I review only a small part of the relevant, but enormous, literature on animal learning. The argument I make is that the training necessary for conducting research on captive animals would not meaningfully affect behavior compared with the contact they normally have in the captive environment. In fact, the impact would probably be less than that resulting from nonresearch interactions with humans. Moreover, the transfer of associations to humans from captive to natural settings is likely to be weak for many behaviors because of contextual influences on memory.

To give this argument proper perspective it is important to describe the types of human contact that exist with releasable animals in captivity outside of any behavioral

research context. I have selected two marine mammals, bottlenose dolphins, a predatory species, and West Indian manatees (*Trichechus manatus*), an herbivorous grazing species, as examples, and because of similarities in learning processes across species, the arguments should apply to other animals. Bottlenose dolphins demonstrate similar associative learning characteristics to other animals (Schusterman 1980). Manatees have been studied less, but initial reports suggest learning consistent with that of other animals (Gerstein et al. 1999; Colbert et al. 2001).

Capture of marine mammals in the United States is restricted by the Marine Mammal Protection Act and Amendments (review in Baur et al. 1999), so dolphins and manatees likely to be released are brought into captivity because of illness, injury, or stranding through rescue programs (Wilkinson & Worthy 1999; U.S. Fish and Wildlife Service 2001). Those animals that survive are rehabilitated and frequently returned to the wild. While in captivity animals have frequent interactions or associations with people during feeding, habitat maintenance, veterinary care, and in some cases public display. They are typically fed by people and/or eat food in the presence of people. What are marine mammals likely to learn in such environments? The answer to this question involves a basic understanding of the core processes of learning (habituation, classical conditioning, and instrumental conditioning, including the concept of stimulus control) and the more complex processes of context-specific memory and its experimental model, compound conditioning. The general principles of learning are briefly reviewed in Griffin et al. (2000) and more extensively described in a variety of texts (e.g., Mackintosh 1974; Dickinson 1980; Pearce 1997; Domjan 2003).

Although not every manatee or dolphin facility follows exactly the same procedures, most share two critical features for learning. The first feature is a frequent exposure of animals to humans (in the absence of explicit research training), which supports habituation. The second is a high correlation of human presence and reinforcement (i.e., food is present and eaten when humans are present, and food is absent and therefore not eaten when humans are absent). If people are present when food is available and not present when food is absent, then the probability increases that people and food will become associated. (Dickinson [1980] and Pearce and Bouton [2001] provide thorough discussions on the development of associations.)

To appreciate more fully the relevance of learning processes to human interactions with captive marine mammals, it is important to understand that in habituation and conditioning, contiguity and covariation among various stimuli and behaviors are important for learning. Correlations between stimuli and behaviors (e.g., people and eating-related behaviors, environments and eating-related behaviors) as well as stimuli and stimuli (e.g., environments and food, people and food, people and pain) strongly influence what is learned.

## Habituation

In a captive situation an animal might initially make various orientation responses toward people or suppress ongoing behaviors in their presence. With repeated exposure to people these behaviors will habituate. Habituation can be defined as a reduced response to repeated stimulation not attributable to fatigue or sensory adaptation (Domjan 2003). It has been studied in a variety of response systems, behavioral and physiological, but the phenomena most relevant to released animals are orientation and suppression responses. No specific behavioral training such as might occur during research procedures is necessary to generate habituation. The regular presence of humans through animal care procedures and viewing by the public and staff will produce it. Exposure to humans in the natural environment apparently leads to habituation in wild dolphins (Lockyer 1990).

Of substantial importance to the release issue is the fact that habituation of orientation and suppression is context dependent (Evans & Hammond 1983; Lovibond et al. 1984; Jordan et al. 2000). When a response habituates in one context, it dishabituates (i.e., returns toward prehabituation levels) in a new context. For example, Peeke and Veno (1973) conducted an experiment in which three-spined sticklebacks (*Gasterosteus aculeatus*) displayed aggressively toward intruding conspecifics. Repeated exposure to the same individual resulted in habituation of display when subjects were tested with the same individual in the same location. Subjects exposed to a new individual in the same location or exposed to the same individual in a new location dishabituated, although not completely (i.e., they resumed aggressive displays, but at a lower rate than the initial level). When exposed to a new fish in a new location, which increased the differences in context, the level of aggressive display returned to or exceeded the original level of response.

In general, whatever habituation of orientation and suppression responses do occur in the captive setting can be expected to dishabituate in the wild because of the substantial differences in context. Furthermore, the phenomenon of spontaneous recovery—the return of a response toward prehabituated levels following the simple passage of time (review in Fantino & Logan 1979)—should further contribute to the attenuation of habituation between a captive and natural environment.

## Classical Conditioning

In classical conditioning a neutral stimulus, the conditioned stimulus (CS), becomes associated with a primary stimulus, the unconditioned stimulus (US), through repeated pairings. For example, in the classic Pavlovian model illustrated in most introductory texts, a biologically significant stimulus, food (US), elicits an unconditioned

response (UR) such as salivation. When an initially neutral stimulus, a bell (CS), is paired with the US, it comes to elicit salivation, the conditioned response (CR). Psychologists have tended to focus on the CS-US relationship in this model. Over the last 30 years some of the most powerful models of learning have been derived from these stimulus-stimulus relationships.

One can use the classical conditioning model to understand what marine mammals learn in the standard free feeding format typically used in captivity. For example, food can be considered a US, and to the degree that a human presence predicts food, it becomes a CS. Hand feeding of foods presents a close temporal-spatial association (contiguity) and correlation between human presence and food consumption. In the case of captive dolphins all feeding is correlated with human presence—this is a particularly strong presence because the food is delivered by humans. Manatees present a slightly less-correlated pattern because they are grazers and large amounts of food are placed in their tanks and are available for eating throughout the day, when humans are not always present. Initial delivery by people is paired with food reward, however, and to the extent that during the day oceanarium viewers and staff are present most of the time, eating is done primarily in the presence of humans. Critically, because food is not made available at night in many facilities, there is an extended period when a “no food, no humans” association is developed. For both dolphins and manatees these feeding patterns mean food and eating occur almost completely in the presence of humans and rarely in their absence. Under such circumstances human presence is predictive of food, a rewarding situation, which learning theory suggests would lead to a strong, excitatory association between humans and food reward (cf. Rescorla 1968).

Training situations present a different pattern of relationships between conditioned stimuli and unconditioned stimuli. In the training situation specific stimuli such as the trainer's whistle or a correctly selected experimental stimulus become associated with food. By pairing the whistle (CS) with food (US), it becomes an effective predictor or substitute for food. Similarly, a rewarded stimulus in a detection or discrimination task becomes associated with food. For example, in a light detection task, the presence of a light becomes associated with food because food is delivered after presentation of a light and is correlated with it. The human trainer is not the predictor of food in these cases; experimental stimuli are. Hence, associations should not develop between humans and food.

## Instrumental Conditioning

Associations are developed between behaviors and stimuli in instrumental conditioning procedures. Animals learn which behaviors are followed by rewards or pun-

ishments and which are not. When rewards (reinforcements) or punishments are only available under specific stimulus conditions, the behavior will be differentially exhibited when these conditions are present. Another way of saying this is that specific, antecedent stimuli called discriminative stimuli ( $S^D$ ) come to determine the performance of a behavior (R, for response). When a behavior is determined by these discriminative stimuli it is said to be under stimulus control. A variety of associations may develop within the instrumental conditioning model, but one that has special importance for understanding my arguments on the effects of training is the stimulus-stimulus association, the association between the discriminative stimulus ( $S^D$ ) and a reinforcing stimulus ( $S^R$ ) such as food. These stimulus-stimulus relationships are essentially classically conditioned associations embedded in the instrumental conditioning framework (Hull 1931; Spence 1956; Rescorla & Solomon 1967).

The delivery of food ( $S^R$ ) in most nontraining interactions at oceanaria is strongly contingent on the presence of humans (i.e., humans are the discriminative stimuli), although depending on reward contingencies items such as food pails or sounds of opening gates may also attain stimulus control. In the research training situations behaviors are brought under the control of specific, experimental discriminative stimuli such as lights, sounds, and trainers' hand signals. Therefore, in the experimental research setting food is not contingent on the mere presence of a person; it results only when a specific behavior is performed in response to a specific discriminative stimulus.

The basic processes influencing an animal's behavior in training circumstances relate to discrimination learning. Subjects have to learn over many trials to discriminate between the specific training stimuli (i.e., experimental stimuli and signals) and the many other irrelevant stimuli, including human-related stimuli. Basically, they come to learn which stimuli predict reward and which do not. This is reflected in increasing numbers of correct responses in the presence of discriminative stimuli that predict reward and decreasing responses to stimuli that do not predict reward. In the behavioral research setting, humans predict reward most frequently when they are signaling and/or when they are accompanied by the paraphernalia associated with experimental research (e.g., targets, manipulanda, audio speakers, and stationing platforms). Unlike the standard, free feeding maintenance condition, humans alone (not signaling or accompanied by research paraphernalia) do not predict reward.

Simple instrumental or classical conditioning, however, is not a fully adequate model to predict the results of more complex human interactions in animal training. Under many research regimens humans are clearly present in conjunction with trainer signals and experimental stimuli. These cases are best considered within the framework of compound conditioning, occasion setting, or contextual effects.

## Compound Conditioning: Elemental and Configural Approaches

Complex context effects can be investigated using a simplified classical conditioning model with a compound CS. For example, humans plus signals or experimental stimuli can be considered compound stimuli, a fact that brings an additional learning process—overshadowing—into play (Rescorla & Wagner 1972; Pearce & Bouton 2001). Overshadowing occurs when one stimulus (CS<sub>1</sub>) interferes with learning about a simultaneously presented stimulus (CS<sub>2</sub>). In general, a more salient stimulus will overshadow a less salient one. For example, within a training procedure humans predict reward at a lower probability level than signals do because when humans are present in the training situation they provide rewards infrequently (or never) when signals are not being given, whereas rewards are provided at a high frequency when a signal (e.g., hand signal, target) is given followed by a correct behavior. Hence signals are more salient than nonsignaling humans. Under a training regimen the subjects learn that the mere presence of humans does not predict reward reliably; only signaling humans predict reward (i.e., learning about humans alone as a predictor of food is overshadowed by learning about signals). The human–food association would be substantially attenuated within this scenario.

Furthermore, under some circumstances overshadowing results in a phenomenon called conditioned inhibition in which the associability of the overshadowed stimulus is actually inhibitory. For example, if humans are out of sensory range during a testing procedure when food reinforcements are provided, then the association between experimental equipment and food will be strong. If humans are then present to remove equipment after completion of a training session when no food is available (i.e., equipment + humans = no food), then humans are likely to form an inhibitory association with food. An inhibitory association is characterized by difficulty in learning a human–food association in the future. Analyzing humans and their signals as separate components of a compound is based on the Rescorla–Wagner model of associative learning (1972), perhaps the most influential theory in learning over the last 30 years. It treats compound stimuli as separable elements, some of which will form excitatory associations with the US, in this case food, and some of which will form inhibitory associations.

Herman et al. (1990) presented an example of the ability of animals to separate manual gestures from the actual human signaler. Two bottlenose dolphins had previously been trained to perform specific behaviors in response to discrete hand signals. The experimenters presented the dolphins with video images of successive degradations of the human hand signals, first by eliminating the head and torso, then the arms, ultimately leaving only images of two flat spots of light moving in black space. Even when

provided with only the spots of light on a video screen, the dolphins were able to interpret the signals correctly.

Testing with successive degradations may have allowed the dolphins to practice separating human gestures from the humans themselves. In a situation that did not entail intentional training, D. Kleiman (personal communication) reports that field assistants carried backpacks containing food, which they distributed throughout the postrelease habitat of golden lion tamarins, and tamarins associated the sound of the backpack zippers with food but did not associate the humans with food. This observation may be explained by the fact that zippers were more reliable predictors of food than humans (i.e., the sound of zippers overshadowed learning about humans).

An influential alternative to the elemental interpretation of learning such as the Rescorla–Wagner approach is the configural model (Pearce 1987). According to this model animals learn about the overall configuration of a compound stimulus rather than the separate elements. Over trials the animal learns the association between a compound CS and a US such as food. If the stimulus compound is altered in some way the associations between CS and US are weakened as reflected in a weaker response. For example, if an animal learns to associate a signaling human with food, then a nonsignaling human will manifest a weaker association because the learned configuration has been altered. In the configural model we predict some initial generalization from signaling human to nonsignaling human based on the similarity of the predictor stimuli. Over time generalization becomes more limited, and the subject clearly discriminates the two different types of stimuli. The implication for training animals is that discrimination between nonsignaling and signaling humans would increase with longer training and generalization would decrease.

Although there is still active discussion among researchers about how learning about stimulus compounds occurs, it is not necessary to analyze that debate here. Sometimes compounds are treated as configural wholes and at others as separable elements (Fanselow 2000; Pearce & Bouton 2001). In either case, the evidence itself and the implications for animal training are clear. Explicit research training of animals should lead to weaker associations between humans and food rewards than that which develops in free-feeding situations in the captive environment. Moreover, under some circumstances associations between nonsignaling humans (the state in which we normally find them) and food are actually inhibited by previous training.

## Compound Conditioning: Modulation

Sometimes an element of a stimulus pair may not form an association with a US, but it does play a role in modulating associations (Holland 1985). In classical conditioning,

modulators are called occasion setters, and they inform the organism that when stimulus A is present stimulus B will be followed by a US. For example, a sound (CS) will predict food (US) when an overhead light is on but not when it is off. In the animal training context, experimental stimuli (CS) predict food (US) when humans are present (occasion setter). If humans are not present, the equipment does not predict food. Within the occasion-setting model a human does not become associated with food but only predicts the CS-US contingency.

The modulator itself does not predict food. It predicts that a stimulus-food or response-food contingency is in effect. This is in sharp contrast to the free feeding situation typically encountered in captive settings where humans become directly associated with food. Or still worse, if human feeders are not careful, they may reinforce a direct approach by providing food when the animal moves toward them. This is a strong learning paradigm in which the human acts as a discriminative stimulus signaling the subject that it will be fed if it approaches the trainer.

## Context-Specific Memory

There is a broader issue than training versus nontraining that affects how one should think about learning in all captive circumstances: the influence of the environment in which a behavior is learned on performance of that behavior in a new environment. Habituation is attenuated in new environments. Why? The answer lies in combining two theories, opponent process theory (Solomon & Corbit 1974; Solomon 1980) and Rescorla-Wagner theory (Rescorla & Wagner 1972).

There is a substantial body of research demonstrating that conditioned responses are not exactly the same as unconditioned responses; in fact, under some circumstances they are the opposite. For example, drug tolerances are frequently mediated by classical conditioned processes in which the physiological response of the organism to a drug is the opposite of that to cues (CSs) predicting the drug (e.g., Siegel 1999). In other words, the CSs set up an opponent process that damps the effect of the drug. A similar situation occurs in the case of habituation. A response is generated by a CS that is opposite to that generated by the US and eventually cancels the response. For example, the orienting response (UR) to a novel object (US) may quickly habituate over multiple exposures because of an opponent CR. But what is the CS?

Rescorla and Wagner (1972) provide an answer to this question by drawing attention to the important role of context in CS-US learning. The Rescorla-Wagner model explains habituation by positing that the environmental context could function as a CS and become associated with the US. In the absence of a specific CS, a US such as a novel object becomes associated with the context. This model provides an explanation for dishabituation in

new contexts. For example, if an animal were to become habituated to a stimulus such as a human presence in a captive context, it would reflect the development of a CS (captive context)-US (human) association. The opponent process CR would damp the orienting response. However, if the CS were not present in opposition to the US, such as would occur in a new environment, then the initial UR, the orienting response, would occur. Occasion setting and other learning processes probably contribute to the role of context as well, but the general conclusion of context specificity remains the same.

Substantial deficits in other types of learning result when animals are tested in environments different from where learning occurred (review in Gordon & Klein 1994). The greater the dissimilarity of environments, the less retention there will be. Interestingly, removing contextual elements reduces transfer but adding elements does not (González et al. 2003).

Context effects are most consistently apparent for inhibitory responses such as extinction (Bouton 1993) in which a previously existing behavior is reduced in frequency. Substantial evidence indicates that changes in context attenuate appetitive (e.g., food rewarded) conditioning (Riccio et al. 1966; Steinman 1967; Chizar & Spear 1969; Rescorla et al. 1985; Hall & Honey 1989; Peck & Bouton 1990). The picture is not, however, entirely consistent on the transfer of appetitive learning between environments. Several researchers have reported no effect of context changes (e.g., Bouton & Peck 1989; Kaye & Mackintosh 1990; Peck & Bouton 1990).

Given some inconsistent data on the effect of context on appetitive conditioning, it is helpful to return to the case studies of appetitive responses of released marine mammals to see what actually occurred under conditions of release. Although most studies of released dolphins and manatees have been insufficiently documented to allow for evaluation of the transfer of learning, these three exceptions provide informative examples of context effects.

Gales and Waples (1993) trained a group of 10 captive- and wild-born Indian Ocean bottlenose dolphins, including a calf and three juveniles, for release from a public display facility where they had lived for up to 10 years. The animals had been trained in both exhibition and husbandry behaviors throughout their captivity, including recall to an underwater signal. Before release they were transferred to a large open-water pen for 3 months, where they were trained to ride the bow and wake of a boat and to approach the underwater recall signal. Despite excellent performance in the pen environment, they did not respond to the underwater signal in the open sea. A few approached the observation boat but not consistently. The lack of response to the underwater signal in the open sea and sporadic approach to an observation boat despite previous food-reinforced training suggest the effects of context change on performance.

In a carefully designed study Wells et al. (1998) provide another example of the lack of transfer between contexts.

They observed and recorded the behavior of two male Atlantic bottlenose dolphins before capture, during 2 years of captivity, and after release. In captivity the subjects were trained using appetitive conditioning for husbandry, behavioral enrichment, and cognitive studies of echolocation. Three to 5.6 years after release, they exhibited no interactions with humans not typically found among wild dolphins and they did not adversely influence social patterns of the host population. The evidence from these two case studies of dolphins supports the argument that dolphins can be trained in captivity without transferring nonadaptive captive learning to the wild.

In another controlled release study, Fellner et al. (2005) used appetitive conditioning procedures to train two Florida manatees in a captive setting to perform a variety of behaviors for food rewards, including approaching a trainer in response to a signal, over a 5-month period. Extinction procedures in the captive setting were then applied to the behaviors (i.e., behaviors that previously had been followed by food reward were no longer rewarded). For administrative reasons the animals were released before extinction was complete. Subsequently, trainers visited the manatees in the field and signaled them to perform the previously trained behaviors. Neither manatee demonstrated any of the captive behaviors in response to signals. Although the extinction procedures cannot be ruled out as contributing to the failure of signals to elicit a response in the field, the strong context dependence of extinction suggests alternative causes. A more likely explanation is that the original training was under tight context control, and the dramatic change in environment from captivity to the wild prevented performance transfer.

## Extinction

There is another important implication of research on compound conditioning and context for public policy. When permits are extended by U.S. agencies for training, extinction of trained behaviors at the end of a study is frequently required before release. This means CSs are presented alone rather than in CS-US pairings. For example, a training whistle, typically preceding food, would be presented without the food US. In instrumental conditioning paradigms, previously rewarded behaviors such as paddle presses are no longer rewarded. As I noted in the discussion of context effects, extinction is strongly context dependent (Bouton 1993). This means that whatever extinction training is done in a captive setting before release is likely to be attenuated by the change to the natural environment.

Of greater concern is the implication of a study by Matzel et al. (1985) that shows that extinguishing the response to an overshadowing stimulus can attenuate overshadowing. If associations with humans are overshadowed in a training situation by experimental stimuli, then

extinguishing the response to those stimuli post-training and, consequently, extinguishing the  $S^D$ - $S^R$  association, will increase the association with humans.

Under those circumstances where positive associations with humans might be expected to persist after release (e.g., open-water training of a dolphin, where the captive and wild environments are similar), aversive conditioning might be a more effective method for discouraging undesirable behavior such as approach to boats after release. Unlike behaviors generated by inhibitory or appetitive processes, fear-related behaviors are resilient to changes in environment (e.g., Bouton & King 1983; Lovibond et al. 1984; Kaye et al. 1987; Hall & Honey 1989). Aversive conditioning, in which undesirable behaviors are followed by a punishing stimulus, would be more likely to discourage orientation toward humans than extinction. The difficulty of appropriate application and collateral effects of punishment such as stress and emotional responding, however, suggest caution in the utilization of aversive techniques.

## Discussion

The clearest way to ensure that animals learn nothing about humans while in captivity is to isolate them completely from any sensory cues of human existence. Such complete isolation, however, is likely to be rare. Captive animals are typically exposed to humans through medical and husbandry procedures, facilities maintenance, and in some cases public display. It would be difficult to totally isolate many species from humans, and not necessarily desirable. Mellen and colleagues (Mellen 1991; Mellen et al. 1998) observed that felids derive notable benefits from interactions with caretakers, including enhanced reproductive success and reduced stress-related behaviors (e.g., pacing). Dierauf (1990) identifies social isolation as a potential risk factor in herd-oriented animals such as many marine mammal species. Providing a stimulating environment also suggests the desirability of research training. Goldblatt (1993), in a review of literature on captive animal stress, concluded that understimulating environments were associated with stress responses in a wide range of animals, including marine mammals. He also concluded that training was the best way to attenuate that stress.

For reasons of practicality and animal welfare, interactions in captivity between many species and humans are likely to remain the norm. As long as animals are going to be in captivity, interacting with humans, it is beneficial to find out something useful for protecting them and their habitats. Many of the characteristics of animals relevant to their conservation, such as what they sense, how they process information, and how they respond physiologically, require behavioral training.

Various researchers have contributed modifications or alternatives to the elemental, configural, and occasion-setting theories I have described (review in Pearce &

Bouton 2001), but they lead essentially to the same conclusion concerning training releasable animals: Associations between humans and pleasurable consequences are less likely to occur in a research-training setting, where animals are brought under stimulus control, compared with other captive interactions such as those associated with free feeding, care, and general viewing. Research on context effects predicts that many of those associations that do develop between humans and pleasurable consequences undergo attenuation when the marine mammals' environments are changed from oceanaria enclosures to natural settings. The notable difference between environments suggests that the attenuation would be substantial. This prediction is supported by the three case studies with marine mammals that have been documented carefully.

It is important to be clear about what is and is not being suggested in my argument. I do not claim that animals learn nothing about humans in behavioral research settings. I suggest that they probably learn no more non-adaptive information about humans than they learn in other circumstances in the captive setting. In some cases research training may attenuate potentially dangerous associations between humans and reward, although it will not always reduce undesirable learning from outside the experimental setting. For example, if people free feed animals, the biological significance of humans as a CS is enhanced considerably. Under such circumstances other CSs such as experimental stimuli may not overshadow humans, even if they are more predictive of reward within the experimental setting. (See Miller and Matute [1996] for a discussion of the effects of biological significance on learning.) This is not a problem of research training; it is a problem of the associations developed outside of research.

It is also important to recognize areas in which the arguments I present may not apply or would at least have to be modified substantially. Training animals in natural settings (e.g., training marine mammals in open water) increases the similarity between training and natural contexts and therefore is more likely to be generalized unless efforts are clearly made to define the research context precisely (i.e., establish tight stimulus control). Lockyer (1990) reviews the case of Dolly, an open-water-trained bottlenose dolphin that was released because of her unpredictable behavior. After release she played with people and allowed them to touch her, behavior ostensibly inconsistent with the arguments for dishabituation and limited transfer of behaviors learned in captivity. Training, however, occurred in the same environment in which they were displayed. In addition, unpredictable behavior by definition indicates a lack of good stimulus control. Therefore it was not surprising that habituation was maintained and behaviors were transferred.

I have not addressed the issue of learning during sensitive periods such as infancy. Animals born and/or reared

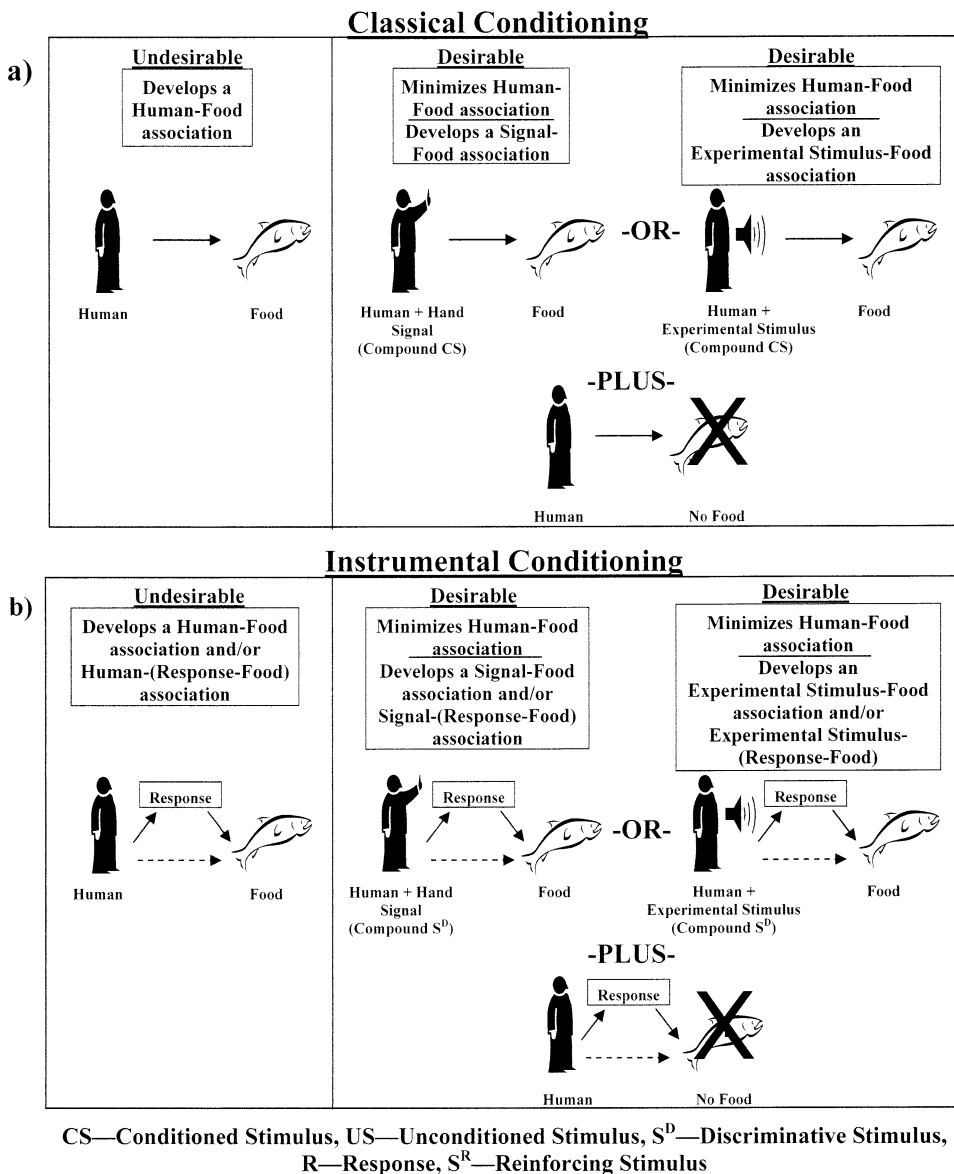
in captivity may form abnormal attachments to people because of the strong learning that sometimes occurs during sensitive, early periods in development. These attachments in conjunction with a lack of normal learning experiences about the natural environment may adversely affect release. This would not, however, be exacerbated by behavioral research.

Within the laboratory setting investigations need to be made on the effects of humans as conditioned or discriminative stimuli. In addition we should conduct carefully controlled experiments to examine the extent to which training of releasable animals in captivity affects their behavior after release. The complex interactions and continuous flow among stimuli and responses in natural environments might generate relationships unpredictable from carefully controlled laboratory studies in which experimental stimuli are frequently discrete and limited in number. Perceptual, motor, motivational, and perhaps higher cognitive factors might interact with basic learning to generate unexpected outcomes. Species and individual characteristics might differ in ways that would affect the salience of key variables. For example, the biological significance of humans may differ among species and certainly will vary depending on individual learning history. The principles of learning are quite stable, although not without some variability (reviews in Shettleworth 1972; Domjan 1983).

Until field experiments can provide direct evidence of training effects, policy concerning human interactions with releasable animals should be based on available empirical evidence. The experimental laboratory evidence suggests that the following practices should be used: (1) Feeding should always be contingent on the presence of distinctive stimuli and animal responses uncorrelated with a human presence. Positive reinforcement uncorrelated with humans minimizes associations between humans and reward. Feeding contingent on human presence alone should be avoided because it conditions animals to associate people with food (Fig. 1). (2) The number of humans interacting with the animals on a noncontingent basis should be limited because it enhances generalization to all humans. (3) Feeding contexts should be made as different from natural contexts as possible. Because removing objects from the learning environment reduces transfer (González et al. 2003), the context should include many different stimuli that will not be present in the natural environment. (4) Extinction may be superfluous because of the behavioral attenuation that would be expected to occur between captive and natural environments, but if it does prove necessary, it should be done in the natural environment. Extinction should also target responses to humans, not to experimental stimuli, because the latter practice might remove overshadowing effects and enhance responses to humans.

Ironically, current practices that limit behavioral research may inadvertently facilitate association of humans





*Figure 1. Training methods for minimizing associations between humans and food. Presenting humans in compounds with other stimuli reduces the association between humans and food. If in addition humans are present when no food is given, the association will be further minimized and under some circumstances may be inhibitory.*

with food, the very characteristic that federal policy is meant to discourage. Animals learn about their environments, including people, with or without explicit training. A critical objective in caring for animals in captivity is that they not learn responses that will transfer to the wild and endanger them. Behavioral training of releasable animals, such as that associated with assessment of sensory processes, cognition, and many types of physiological research, provides an excellent solution to the problem of minimizing undesirable associations with people, providing environmental enrichment, and adding knowledge of species important for their conservation.

## Acknowledgments

This manuscript benefited greatly from the perceptive comments of D. E. Colbert, W. Fellner, H. E. Harley, D.

A. Mann, S. Partan, R. Reep, J. E. Reynolds III, H. L. Roitblat, R. S. Wells, J. Mellen, Assigning Editor D. Kleiman, and three anonymous reviewers. Joseph C. Gaspard III prepared the figure. N. Adimey and J. Whaley were very helpful in providing sources that clarified the policy of U.S. Fish and Wildlife Service and National Marine Fisheries Service concerning marine mammal training and research. Grants from the Florida Fish and Wildlife Commission contributed to the preparation of this manuscript.

## Literature Cited

- Baur, D. C., M. J. Bean, and M. L. Gosliner. 1999. The laws governing marine mammal conservation in the United States. Pages 48–86 in J. R. Twiss Jr. and R. R. Reeves, editors. Conservation and management of marine mammals. Smithsonian Institution Press, Washington, D.C.
- Bouton, M. E. 1993. Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin* 114:80–99.

- Bouton, M. E., and D. A. King. 1983. Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes* 9:248-265.
- Bouton, M. E., and C. A. Peck. 1989. Context effects on conditioning extinction and reinstatement in an appetitive conditioning preparation. *Animal Learning & Behavior* 17:188-198.
- Chizar, D. A., and N. E. Spear. 1969. Stimulus change, reversal learning, and retention in the rat. *Journal of Comparative and Physiological Psychology* 69:190-195.
- Colbert, D. E., W. Fellner, G. B. Bauer, C. A. Manire, and H. L. Rhinehart. 2001. Husbandry and research training of two Florida manatees (*Trichechus manatus longirostris*). *Aquatic Mammals* 27:16-23.
- Dickinson, A. 1980. Contemporary animal learning theory. Cambridge University Press, Cambridge, United Kingdom.
- Dierauf, L. A. 1990. Stress in marine mammals. Pages 295-301 in L. A. Dierauf and J. Mazzeo, editors. CRC handbook of marine mammal medicine. CRC Press, Boca Raton, Florida.
- Domjan, M. 1983. Biological constraints on instrumental and classical conditioning: implications for general process theory. *Psychology of Learning and Motivation* 17:215-277.
- Domjan, M. 2003. The principles of learning and behavior. 5th edition. Wadsworth/Thomson Learning, Belmont, California.
- Evans, J. G. M., and G. R. Hammond. 1983. Differential generalization of habituation across contexts as a function of stimulus significance. *Animal Learning and Behavior* 11:432-434.
- Fanselow, M. S. 2000. Learning theory and neuropsychology: configuring their disparate elements in the hippocampus. *Journal of Experimental Psychology: Animal Behavior Processes* 25:275-283.
- Fantino, E., and C. A. Logan. 1979. The experimental analysis of behavior: a biological perspective. W. H. Freeman and Company, San Francisco, California.
- Fellner, W., K. Odell, A. Corwin, L. Davis, C. Goonen, I. Larkin, and A. M. Stamper. 2005. Response to conditioned stimuli of two released rehabilitated West Indian manatees (*Trichechus latirostris*). *Aquatic Mammals* (in press).
- Gales, N., and K. Waples. 1993. The rehabilitation and release of bottlenose dolphins from Atlantic Marine Park, Western Australia. *Aquatic Mammals* 19:49-59.
- Gerstein, E., L. Gerstein, S. Forsythe, and J. Blue. 1999. The underwater audiogram of the West Indian manatee (*Trichechus manatus*). *Journal of the Acoustical Society of America* 105:3575-3583.
- Goldblatt, A. 1993. Behavioral needs of captive marine mammals. *Aquatic Mammals* 19:149-157.
- González, E., J. J. Quinn, and M. S. Fanselow. 2003. Differential effects of adding and removing components of a context on the generalization of conditional freezing. *Journal of Experimental Psychology: Animal Behavior Processes* 29:78-83.
- Gordon, W. C., and R. L. Klein. 1994. Animal memory: the effects of context change on retention performance. Pages 255-279 in N. J. Mackintosh, editor. *Animal learning and cognition*. Academic Press, San Diego, California.
- Griffin, A. S., D. T. Blumstein, and C. S. Evans. 2000. Training captive-bred or translocated animals to avoid predators. *Conservation Biology* 14:1317-1326.
- Hall, G., and R. C. Honey. 1989. Contextual effects in conditioning, latent inhibition, and habituation: associative and retrieval functions of contextual cues. *Journal of Experimental Psychology: Animal Behavior Processes* 15:232-241.
- Herman, L. M., P. Morrel-Samuels, and A. A. Pack. 1990. Bottlenose dolphin and human recognition of veridical and degraded video displays of an artificial gestural language. *Journal of Experimental Psychology: General* 119:215-230.
- Holland, P. C. 1985. The nature of conditioned inhibition in serial and simultaneous feature negative discriminations. Pages 267-297 in R. Miller and N. E. Spear, editors. *Information processing in animals: conditioned inhibition*. Erlbaum, Hillsdale, New Jersey.
- Hull, C. L. 1931. Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review* 38:487-506.
- Jordan, W. P., H. C. Strasser, and L. McHale. 2000. Contextual control of long-term habituation in rats. *Journal of Experimental Psychology: Animal Behavior Processes* 26:323-339.
- Kleiman, D. G. 1989. Reintroduction of captive mammals for conservation. *Bioscience* 39:152-161.
- Kaye, H., and N. J. Mackintosh. 1990. A change of context can enhance performance of an aversive but not of an appetitive conditioned response. *Quarterly Journal of Experimental Psychology* 42B:113-134.
- Kaye, H., G. C. Preston, L. Szabo, H. Druiff, and N. J. Mackintosh. 1987. Context specificity of conditioning and latent inhibition: evidence for a dissociation of latent inhibition, and associative interference. *Quarterly Journal of Experimental Psychology* 39B:127-145.
- Lockyer, C. 1990. Review of incidents involving wild, sociable dolphins, worldwide. Pages 337-353 in S. Leatherwood and R. R. Reeves, editors. *The bottlenose dolphin*. Academic Press, San Diego, California.
- Lovibond, P. F., G. C. Preston, and N. J. Mackintosh. 1984. Context specificity of conditioning, extinction, and latent inhibition. *Journal of Experimental Psychology: Animal Behavior Processes* 10:360-375.
- Mackintosh, N. J. 1974. *The psychology of animal learning*. Academic Press, London, United Kingdom.
- Macphail, E. M. 1982. *Brain and intelligence in vertebrates*. Clarendon Press, Oxford, United Kingdom.
- Manire, C., C. J. Walsh, H. L. Rhinehart, D. E. Colbert, D. R. Noyes, and C. A. Luer. 2003. Alterations in blood and urine parameters in two Florida manatees (*Trichechus manatus latirostris*) from simulated conditions of release following rehabilitation. *Zoo Biology* 22:103-120.
- Matzel, L. D., T. R. Schachtman, and R. R. Miller. 1985. Learned irrelevance exceeds the sum of CS-preexposure and US-preexposure deficits. *Journal of Experimental Psychology: Animal Behavior Processes* 14:311-319.
- Mellen, J. 1991. Factors influencing reproductive success in small captive exotic felids (*Felis* spp.): a multiple regression analysis. *Zoo Biology* 10:95-110.
- Mellen, J. D., M. P. Hayes, and D. J. Shepherdson. 1998. Captive environments for small felids. Pages 184-201 in D. J. Shepherdson, J. D. Mellen, and M. Hutchins, editors. *Second nature: environmental enrichment for captive animals*. Smithsonian Institution Press, Washington, D.C.
- Miller, R. R., and H. Matute. 1996. Biological significance in forward and backward blocking: resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General* 125:370-386.
- National Oceanic and Atmospheric Administration (NOAA). 2003. U.S. Code of Federal Regulations (CFR), 50 CFR, Section 216.27. U.S. Government Printing Office, Washington, D.C.
- Papini, M. R. 2002. Patterns and process in the evolution of learning. *Psychological Review* 109:186-201.
- Pearce, J. M. 1987. A model of stimulus generalization in Pavlovian conditioning. *Psychological Review* 94:61-73.
- Pearce, J. M. 1997. *Animal learning and cognition*. Psychology Press, East Sussex, United Kingdom.
- Pearce, J. M., and M. E. Bouton. 2001. Theories of associative learning in animals. *Annual Review of Psychology* 52:111-139.
- Peck, C. A., and M. E. Bouton. 1990. Context and performance in aversive-to-appetitive and appetitive-to-aversive transfer. *Learning and Motivation* 21:1-31.
- Peeke, H. V. S., and A. Veno. 1973. Stimulus specificity of habituated aggression in three-spined sticklebacks (*Gasterosteus aculeatus*). *Behavioral Biology* 8:427-432.

- Rescorla, R. A. 1968. Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology* **66**:1-5.
- Rescorla, R. A., and R. L. Solomon. 1967. Two-process theory: relationship between Pavlovian conditioning and instrumental learning. *Psychological Review* **88**:151-182.
- Rescorla, R. A., and A. R. Wagner. 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. Pages 64-99 in A. H. Black and W. F. Prokasy, editors. *Classical conditioning II: current research and theory*. Appleton-Century-Crofts, New York.
- Rescorla, R. A., P. J. Durlach, and J. W. Grau. 1985. Contextual learning in Pavlovian conditioning. Pages 23-56 in P. D. Balsam and A. Tomie, editors. *Context and learning*. Earlbaum, Hillsdale, New Jersey.
- Riccio, D. C., M. Urdu, and D. R. Thomas. 1966. Stimulus control in pigeons based on proprioceptive stimuli from the floor inclination. *Science* **153**:434-436.
- Shettleworth, S. J. 1972. Constraints on learning. Volume 4. Pages 1-68 in D. S. Lehrman, R. A. Hinde, and E. Shaw, editors. *Advances in the study of behavior*. Academic Press, New York.
- Schusterman, R. J. 1980. Behavioral methodology in echolocation by marine mammals. Pages 11-41 in G. Busnel and J. R. Fish, editors. *Animal sonar systems*. Plenum Press, New York.
- Siegel, S. 1999. Drug anticipation and drug addiction. *Addiction* **84**: 1113-1124.
- Solomon, R. L., and J. D. Corbit. 1974. An opponent-process theory of motivation. I. The temporal dynamics of affect. *Psychological Review* **81**:119-145.
- Solomon, R. S. 1980. The opponent process theory of acquired motivation. *American Psychologist* **35**:691-712.
- Spence, K. W. 1956. *Behavior theory and conditioning*. Yale University Press, New Haven, Connecticut.
- Steinman, F. 1967. Retention of alley brightness in the rat. *Journal of Comparative and Physiological Psychology* **64**:105-109.
- U.S. Fish and Wildlife Service. 2001. Florida manatee recovery plan (*Trichechus manatus latirostris*). Third revision. U.S. Fish and Wildlife Service, Atlanta.
- U.S. National Marine Fisheries Service (NMFS) and U.S. Fish and Wildlife Service (USFWS). 1997. Draft release of stranded marine mammals to the wild: background, preparation, and release criteria. NMFS-OPR technical memorandum. National Oceanic and Atmospheric Administration, Washington, D.C. Available from [http://www.nmfs.noaa.gov/prot\\_res/readingrm/MMHealth/mmreleaseguid.pdf](http://www.nmfs.noaa.gov/prot_res/readingrm/MMHealth/mmreleaseguid.pdf) (accessed November 2004).
- Wells, R. S., K. Bassos-Hull, and K. S. Norris. 1998. Experimental return to the wild of two bottlenose dolphins. *Marine Mammal Science* **14**:51-71.
- Wilkinson, D., and G. A. J. Worthy. 1999. Marine mammal stranding networks. Pages 396-411 in J. R. Twiss Jr. and R. R. Reeves, editors. *Conservation and management of marine mammals*. Smithsonian Institution Press, Washington, D.C.

